J. theor. Biol. (1989) **140,** 453–477

A Model for the Initiation of Replication in *Escherichia coli*

JOSEPH M. MAHAFFY†

Department of Mathematical Sciences, San Diego State University, San Diego, CA 92182, *U.S.A.*

AND

JUDITH W. ZYSKIND‡

Department of Biology and Molecular Biology Institute, San Diego State University, San Diego, CA 92182, *U.S.A.*

(*Received* 7 *February* 1989, *and accepted in revised form* 15 *May* 1989)

The role of the protein DnaA as the principal control of replication initiation is investigated by a mathematical model. Data showing that DnaA is growth rate regulated suggest that its concentration alone is not the only factor determining the timing of initiation. A mathematical model with stochastic and deterministic components is constructed from known experimental evidence and subdivides the total pool of DnaA protein into four forms. The active form, DnaA · ATP, can be bound to the origin of replication, *oriC*, where it is assumed that a critical level of these bound molecules is needed to initiate replication. The active form can also exist in a reserve pool bound to the chromosome or a free pool in the cytoplasm. Finally, a large inactive pool of DnaA protein completes the state variables and provides an explanation for how the DnaA · ATP form could be the principal controlling element in the timing of initiation. The fact that DnaA protein is an autorepressor is used to derive its synthesis rate. The model studies a single exponentially growing cell through a series of cell divisions. Computer simulations are performed, and the results compare favorably to data for different cell cycle times. The model shows synchrony of initiation events in agreement with experimental results.

1. Introduction

Initiation of DNA replication begins at a unique sequence called the origin of replication or *oriC*. The nucleotide sequence of *oriC* is known for a number of enteric bacteria (Zyskind *et al.*, 1983), and within *oriC* there are four highly conserved sets of sequences, R1, R2, R3, and R4, that have the consensus TTAT(C/A)CA(C/A)A and serve as DnaA protein binding sites (Fuller *et al.*, 1984; Matsui *et al.*, 1985). DnaA protein is required in order for initiation from *oriC* to occur, as demonstrated by genetic and physiological approaches (Zyskind *et al.*, 1977; Kogoma & von Meyenburg, 1983; Hansen *et al.*, 1984; Atlung, 1984; Skarstad *et al.*, 1988) as well as by use of cell-free and purified *oriC*-dependent replication systems (Fuller & Kornberg, 1983; Ogawa *et al.*, 1985; Funnel *et al.*, 1986; Bramhill & Kornberg, 1988). Only DnaA protein associated with ATP is active in supporting

† This work was supported in part by NSF grant number DMS-8807360.
‡ This work was supported in part by NSF grant numbers DMB-8611081 and DMB-8911672.

0022-5193/89/200453 + 25 $03.00/0

replication from *oriC* in the purified replication systems (Sekimizu *et al.*, 1987). Forms of DnaA protein either unbound to a nucleotide, bound to ADP, or bound to acidic phospholipids, have no activity (Sekimizu & Kornberg, 1988). The first step in initiating *in vitro* replication from *oriC* is the binding of 20-40 monomers of DnaA protein to *oriC* by affinity for the R repeats in *oriC* (Fuller *et al.*, 1984; Funnell *et al.*, 1987). An open complex, which contains an opening of the DNA double helix, is then formed. Proteins involved with priming of replication bind to the complex, and an RNA primer is synthesized that is subsequently extended by DNA polymerase III holoenzyme (Bramhill & Kornberg, 1988).

The concentration of DnaA protein appears to be a critical factor in determining the timing of initiation in the cell cycle (Atlung *et al.*, 1987; Løbner-Olesen *et al.*, 1989). In the cell, the concentration of a protein is determined by the rates of synthesis and degradation and by dilution resulting from cell growth. DnaA protein appears to be a stable protein in that there is little decrease in the concentration after 1 hr of treatment with chloramphenicol (Chiaramello, personal observations); thus, the concentration of DnaA protein will depend mainly on its synthesis rate and dilution during cell growth.

DnaA protein is autoregulated in that the synthesis of DnaA protein is controlled by two promoters, both of which are repressed by DnaA protein (Braun *et al.*, 1985; Atlung *et al.*, 1985; Kücherer *et al.*, 1986; Wang & Kaguni, 1987; Nozaki *et al.*, 1988). DnaA protein inhibits transcription from these two promoters by binding to the DnaA binding site between them. DnaA protein also appears to cause termination of transcription at the intragenic DnaA binding site (Messer *et al.*, 1988). It is not known whether all forms of DnaA protein bind with equal affinity or whether binding of the different forms is equally effective in repression of transcription from the *dnaA* promoters.

In section 2, experimental data for cells grown at different growth rates are used to show that a simple model which depends only on the total concentration of DnaA fails to predict that DnaA protein is the controlling element for initiation of replication. In section 3, a more complicated mathematical model is presented which shows that DnaA protein could be the principal component regulating the initiation of replication. This model incorporates many details that are known about DnaA protein and *oriC* and proposes several avenues for further experimental studies. In our model there are five distinct states for DnaA. Figure 1 shows a schematic for the proposed model.

(1) *Activator.* DnaA when complexed with ATP (DnaA · ATP) can be bound to *oriC* where it serves as an activator for initiation of replication (Sekimizu *et al.*, 1987).

(2) *DnaA · ATP bound form.* It can be bound to any number of gene sites along the chromosome. These sites act as a reserve for DnaA · ATP.

(3) *Autorepressor.* As stated above, DnaA protein acts as a repressor for the expression of its own gene, *dnaA.*

(4) *DnaA · ATP (free).* The fourth form is DnaA · ATP, which is free in the cytoplasm to bind to any of the above gene locations including *oriC.*

(5) *$DnaA_i$ (inactive).* It can also exist in inactive forms, DnaA unbound to a nucleotide, DnaA · ADP, and DnaA bound to acidic phospholipids (Sekimizu &

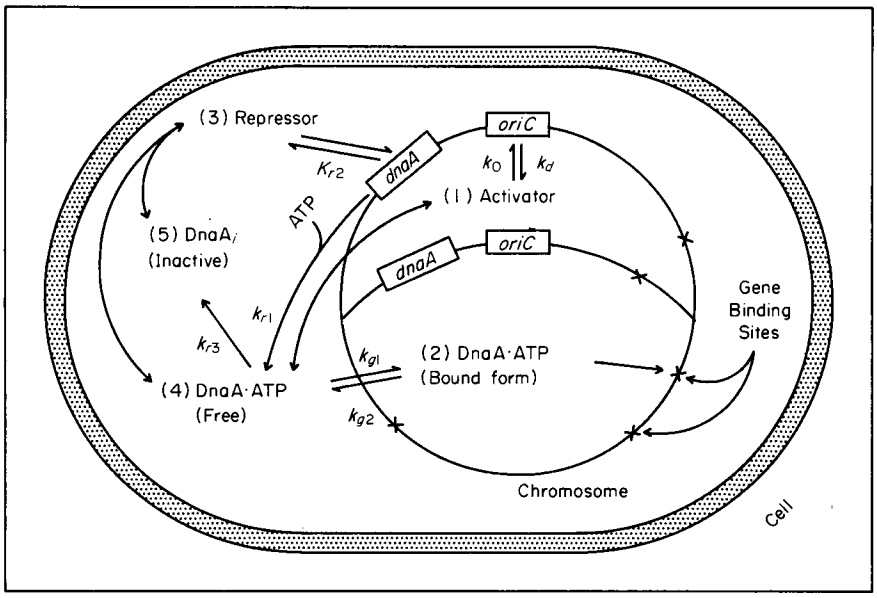

FIG. 1. A schematic for the model.

Kornberg, 1988). We will denote all inactive forms by $DnaA_i$. These five distinct states are interrelated by biochemical reactions. Computer simulations are performed for a series of growing cells using these reactions, and the results are compared to the experimental data.

Several theoretical models have been proposed to describe the control of timing of initiation. The Replicon Hypothesis (Jacob *et al.*, 1963) was one of the first such models. In this model, the frequency and timing of initiation are determined by interaction between an initiator protein and a fixed activator site on each replicon, and membrane attachment of replicons insures correct segregation into daughter cells. A recent version of this model has been proposed by Ogden *et al.* (1988), which involves the binding of hemimethylated *oriC* DNA to the membrane. They demonstrated that there is an 8–10 min period after initiation of replication when hemimethylated *oriC* DNA is bound to membrane material, and proposed that this could be important in separating the replicating chromosomes at division. The Autorepressor Model (Sompayrac & Maaløe, 1973) proposes that an autorepressor and initiator are cotranscribed, and the synthesis of the initiator is regulated by the autorepressor. Initiation occurs when a certain concentration of initiator molecules is reached. Applying this model to the *oriC* system, DnaA would function as both initiator and autorepressor. Our model is based on both the Replicon Hypothesis as now formulated by Ogden *et al.* (1988) and the Autorepressor Model. Our techniques for modeling DNA replication of the chromosome use kinetic equations similar to those developed for the control of plasmid replication (Womble & Rownd, 1987). Our model is being proposed to provide an explanation for some of the

observations made so far with regard to the initiation process and to suggest important parameters for which more information would be most useful in the testing and development of this model.

2. DnaA Protein and *oriC*

One of the most important parameters in our model is the DnaA protein concentration in the cell. This has been determined by Chiaramello & Zyskind (1989) for several different growth rates using immunoblot analysis, and some of the data from that work is shown in Table 1. Although DnaA protein is autoregulated, it was found that the concentration of DnaA protein in the cell, which is reflected in the DnaA/mass and DnaA/protein ratios, increased with growth rate rather than remaining constant, suggesting that further regulatory elements are involved in controlling the expression of the *dnaA* gene and/or that the different forms of DnaA protein are not equally effective in repressing expression of *dnaA*.

Donachie (1968) hypothesized that initiation depended on the attainment of a critical cell mass, called the initiation mass, after observing that the cell mass per origin at the time of initiation was independent of growth rate. An explanation for this correlation between cell mass and the number of chromosomal origins at the time of initiation is that initiation depends on the accumulation of a critical amount of an initiator whose accumulation is correlated to total protein synthesis and, therefore, is proportional to an increase in cell mass. To test whether the amount of DnaA protein/origin at the time of initiation remained constant, we estimated when initiation occurred in the cell cycle and the number of origins of DNA replication in the cell at the time of initiation for the cultures used in the study performed by Chiaramello & Zyskind (1989). Previously estimated values (Bremer & Dennis, 1987; Helmstetter, 1987) for C, the time for a round of chromosomal replication, and D, the time between completion of chromosome replication and cell division, were used.

Except for the slowest growing culture, the initiation mass (mass/origin at initiation, Table 1) remained constant at different growth rates in agreement with the observation of Donachie (1968). One interpretation of the constancy in initiation mass is the presence of a critical protein whose concentration determines when initiation events occur in the cell cycle as mentioned above. The amount of protein per origin (Bremer & Dennis, 1987) and protein per origin at initiation (Table 1, Fig. 2) remained constant at different growth rates, except for the slowest growing cells in agreement with this interpretation. DnaA protein concentration per origin at the time of initiation, however, varied with growth rate, exhibiting a 6·7-fold increase (from 55 to 369 monomers/origin at initiation) as the doublings per hr increased from 0·55 to 2·0 (Table 1, Fig. 2). If different times for C and D were used instead of the C and D values given in Table 1, the results were qualitatively the same. For example, C and D values of 40 min and 20 min, respectively, at all growth rates except for one, the amount of DnaA protein per origin at initiation was approximately 20% higher (data not shown). In conclusion, the number of DnaA protein monomers per origin at the time of initiation varies directly with

TABLE 1

DnaA protein in cells grown at different growth rates

Parameter	τ (doubling time, min) and μ (doublings/hr)					
	τ109, μ0·55	τ75, μ0·80	τ58, μ0·97	τ45, μ1·33	τ38, μ1·58	τ30, μ2·00
Cells/mass (10^8 cells/OD_{450})[a]	12·89	5·83	5·10	3·72	3·17	2·27
Mass (10^{-9} OD_{450})/cell[a]	0·776	1·715	1·961	2·688	3·155	4·405
Mass (10^{-9} OD_{450})/newborn cell[b]	0·55	1·21	1·38	1·89	2·22	3·10
$O(i)$, number of origins/cell at initiation[c]	1·0	1·0	2·0	2·0	2·0	4·0
Mass (10^{-10} OD_{450})/origin at initiation[d]	5·79	12·08	11·96	12·88	12·40	12·31
Protein (10^7 aa)/cell[a]	39·7	91·5	77·2	118·7	135·3	187·1
Protein (10^7 aa)/origin at initiation[d]	29·6	64·5	47·1	56·9	53·2	52·3
% DnaA protein of total protein	0·006	0·005	0·009	0·014	0·016	0·020
DnaA (10^9 monomers)/mass (OD_{450})[a]	95	84	119	215	255	300
DnaA monomers/cell[a]	74	144	234	579	803	1319
DnaA monomers/newborn cell[b]	60	101	165	407	566	929
DnaA monomers/origin at initiation[d]	55	101	143	277	316	369
dnaA genes/cell[e]	1·86	1·97	2·27	2·88	3·50	4·88
Chromosomes/cell[f]	1·5	1·6	1·8	2·2	2·5	3·1
DnaA monomers/*dnaA* gene[g]	40	73	103	201	230	270
DnaA monomers/chromosomal equivalent[h]	49	90	130	263	321	426

[a] Data from Chiaramello & Zyskind (1989).

[b] The concentration of a population parameter in the newborn cell, P_0, is calculated from the age distribution function of Powell (1956).

[c] The number of orgins at the time of initiation varies discontinuously with growth rate (Cooper & Helmstetter, 1968) and was determined from C, D, and τ (Bremer & Dennis, 1987; Helmstetter, 1987). The C period is the period of time in the cell cycle occupied by DNA replication, and the D period is the time between termination of DNA replication and cell division. The C and D values were not determined for the *E. coli* K-12 strain used in the study of Chiaramello & Zyskind (1989), but because they are relatively constant for *E. coli* B/r and K-12 strains, estimates for these values were chosen from previous determinations (see Table 1 of Bremer & Dennis, 1987; Table 1 of Helmstetter, 1987). A value of 25 min for D was used for all growth rates; values for C were 75 min ($\tau = 109$), 50 min ($\tau = 75$), and 45 min ($\tau = 58$, 45, 38 and 30 min).

[d] The cells are assumed to be growing exponentially implying that a population or mass variable, $P(t)$, satisfies the differential equation, $dP/dt = kP$, $P(0) = P_0$ where k is the specific rate of growth, μ is the doublings per hr, $\mu = k/\ln 2$, $\tau = 60/\mu$, P_0 is the initial value of the variable, P. Solution of this initial value problem (Powell, 1956; Ingraham *et al.*, 1983) yields the equation $P(i) = P_0 \cdot 2^{i/\tau}$, where $P(i)$ is a population parameter such as cell mass, protein, or DnaA monomers per cell at initiation, and P_0 is the cell mass, protein, or DnaA monomers per newborn cell. The time in the cell cycle when initiation occurs (i) is equal to $[1+n]\tau - [C+D]$ where n is the smallest integer so that $(n+1)\tau \geq C+D$ (Cooper & Helmstetter, 1968). The concentration of the population parameter at the time of initiation, $P(i)$ is then divided by the number of origins at the time of initiation, $O(i)$, to determine the amount of the parameter relative to the number of origins in the cell at the time of initiation. Nearly identical values were obtained using the relationship, parameter per origin in single cells at initiation, $P(i)/O(i)$, equals the amount of the parameter per origin in an exponential culture divided by ln 2 (Bremer *et al.*, 1979).

[e] The mean number of copies of a gene per cell ($\bar{X}$) in an exponential phase culture is $\bar{X} = 2^{[(1-x)C+D]/\tau}$, where x is the fraction of the C period at which the gene replicates (Helmstetter, 1987). The *dnaA* gene is located at 83 min on the *E. coli* map (Bachmann, 1987), so $x = 0{\cdot}02$.

[f] Average number of chromosomal equivalents (mass of DNA corresponding to a single nonreplicating chromosome) of DNA per cell was calculated from the following formula: $\bar{G} = (\tau/C \ln 2)[2^{(C+D)/\tau} - 2^{D/\tau}]$ (Helmstetter, 1987).

[g] DnaA monomers/cell divided by *dnaA* genes/cell.

[h] DnaA monomers/cell divided by chromosomal equivalents/cell.

growth rate, eliminating the possibility that the total concentration of DnaA is the sole determinant of when initiation occurs.

DnaA protein is autoregulated, and it is predicted that the concentration of an autoregulated protein will not vary with growth rate but remain constant at different growth rates (Sompayrac & Maaløe, 1973); however, as described above, the ratio of DnaA protein to total protein increased with growth rate, indicating that DnaA

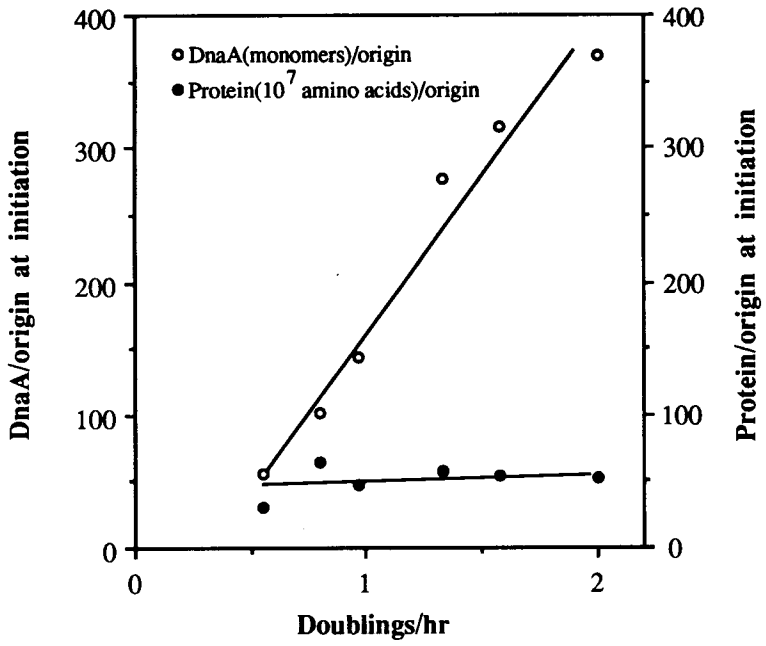

FIG. 2. DnaA protein concentration relative to number of origins at initiation as a function of growth rate (data from Table 1).

protein is growth rate-regulated. One possible unifying mechanism for autoregulation and growth rate-dependent regulation of *dnaA* gene expression involves titration of the DnaA protein by increasing the numbers of DnaA binding sites. DnaA protein binds to DNA fragments containing the sequence TTATCCACA and to variations of this sequence (Fuller *et al.*, 1984). This binding site is found in the promoter regions of many genes and is repeated four times in the origin of replication, *oriC* (Fuller *et al.*, 1984; Matsui *et al.*, 1985). DnaA protein binding to *oriC* results in the formation of large spherical complexes containing 20–40 monomers of DnaA (Fuller *et al.*, 1984; Funnell *et al.*, 1987), so the number of DnaA monomers bound to the sites around the chromosome may be greater than the number of DnaA binding sites. At the slower growth rates, the number of DnaA monomers per cell is sufficiently low to suggest that most of the DnaA is bound to DNA rather than free in the cell. It is unlikely that titration of DnaA protein by DNA binding is responsible for the growth rate-regulated expression of DnaA protein because, if the number of DnaA binding sites is scattered randomly throughout the chromosome, the number of DnaA binding sites in the cell should be proportional to the number of chromosomes in the cell, and there is an 8·7-fold increase in the number of DnaA monomers per chromosome as τ increases from 0·55 to 2·0 doublings per hr (Table 1). Furthermore, because the number of DnaA protein monomers per *dnaA* gene increases 6·7-fold over the range of growth rates examined (Table 1), it is unlikely

that gene dosage of the *dnaA* promoter is responsible for the growth rate-regulated expression.

3. The Mathematical Model

In this section a mathematical model for the initiation of DNA replication in *E. coli* is described. A detailed description of the kinetic equations and the relevant parameters are included in Appendices A and B. The model centers on the role of the protein, DnaA, in the control of replication of *E. coli*. The mathematical model employs many of the experimental results discussed in the previous sections. In addition, there are several assumptions that are needed to formulate a detailed model. These assumptions are described in this section with a discussion of their biological implications to the cell.

The mathematical model uses the biochemical interactions between the five states of DnaA protein illustrated in Fig. 1. The model contains a stochastic process for simulating the binding of DnaA · ATP to *oriC*, while the remainder of the model is considered to be deterministic. The model presents a plausible explanation for the data shown in Table 1 by examining an individual exponentially growing cell through a specified period of time. The computer simulation follows the exponentially growing cell through a series of initiations and cell divisions. We note that only one cell is followed after the cell divides.

A stochastic process is used to compute the amount of DnaA · ATP bound to *oriC*, the site of initiation of DNA replication. In Fig. 1, this DnaA · ATP is designated *activator*. As noted in the introduction, there are four DnaA binding sites in *oriC*. For modeling purposes it is assumed that DnaA · ATP molecules bind to the four binding sites in *oriC* with equally likely chances. In addition, experiments show that DnaA protein binds to itself to form large spherical complexes containing *oriC* DNA and 20–40 monomers of DnaA protein (Fuller *et al.*, 1984; Funnell *et al.*, 1987). In the model, additional molecules are added at the four sites with equally likely random events until an initiation event occurs. For computational purposes, we assume an average rate of binding, k_0, and that there are always four binding sites available for binding of DnaA · ATP. Our assumptions ignore the probable differences in binding constants and changes in the geometry as multiple molecules bind to *oriC*. The model could adapt to either a different number of binding sites by changing the average binding constant or the use of different binding constants which depend on the number of DnaA · ATP monomers which are bound to *oriC*; however, there are insufficient data to justify adding these complications to the model.

The model also considers the possibility that a molecule of DnaA · ATP may dissociate over a small interval of time with dissociation constant, k_d. We assume that only the upper most molecules in the complex are subject to dissociation, which allows a maximum of four molecules corresponding to the four binding sites to be subjected to this process. Future studies are planned to investigate the importance of state dependence of k_0 and k_d to the model.

Before we examine the stochastic binding of DnaA to *oriC* in our model, we consider how long it has been since the last initiation event for a particular *oriC*.

It is known that after replication begins, the chromosomal origin is in a hemimethylated state for 8–10 min during which time it is bound to membranes (Ogden *et al.*, 1988). In the cell, the binding of DnaA protein to *oriC* just after initiation may be prevented by the association of *oriC* with the membrane (Ogden *et al.*, 1988) and by loss of supercoiling (Inman & Schnös, 1987). Affinity of DnaA protein for *oriC* is decreased for non-supercoiled DNA (Fuller & Kornberg, 1983). This altered state is believed to affect the binding ability of certain molecules. The model assumes that for 8 min following an initiation event, DnaA · ATP is unable to bind to *oriC*. If the last initiation event for a particular *oriC* has occurred less than 8 min earlier, then the model does not allow binding and the subsequent probabilistic model does not apply to this particular *oriC*.

After this period of latency for a particular *oriC*, the net probabilities for addition or removal of DnaA · ATP are computed, and a Monte Carlo simulation scheme is used to update the number of DnaA · ATP monomers bound to *oriC* over a small interval Δt. As $\Delta t = 0{\cdot}01$ min in our simulation, which would represent too few events for the time interval, the model averaged R random events as described above with $R = 5$, 10 or 20. This procedure was repeated 10 times for each *oriC* in the growing cell to make a time step of 0·1 min, and then the simulation enters the deterministic part of the model.

The mathematical model assumes that there is a uniform distribution of the consensus sequence along the chromosome where DnaA · ATP can bind. This serves as a reserve for the active form, DnaA · ATP. With the assumption that these binding sites are distributed uniformly along the chromosome, the model adds binding sites regularly to each elongating chromosome. The number of actively elongating chromosomes is determined by the number of *oriC*'s which have an age between 0 and C, where C is the time required for replication of the chromosome. For the cell generation times used in our simulations, C was taken to be 45 min, based on previous observations (Bremer & Dennis, 1987; Helmstetter, 1987).

The next stage in the mathematical model is to determine the concentrations of $DnaA_g$ which are the DnaA · ATP monomers bound to the consensus sequences away from *oriC*, $DnaA_f$ which are the DnaA · ATP monomers that are free in the cytoplasm, and $DnaA_i$ which is the inactive pool of DnaA monomers. These concentrations are governed by non-linear differential equations that are derived using standard biochemical interactions and solved by numerical integration. The equation for the concentration of $DnaA_g$, $[DnaA_g]$, is derived by applying mass balance laws for the formation and release of DnaA · ATP monomers bound to the consensus sequences along the chromosome and including a term to account for dilution. The model does not consider the low intrinsic activity that might convert part of the bound form to DnaA · ADP (Sekimizu *et al.*, 1988).

From Fig. 1, the third state of DnaA in the mathematical model is that of an autorepressor for the *dnaA* gene. We assume that the repressor concentration is determined from the concentration of DnaA · ATP monomers in the cytoplasm and from certain forms of DnaA that are in the inactive state. We use this information and standard biochemical kinetics (Goodwin, 1965; Othmer, 1976) to find the nonlinear production term for the differential equation describing $[DnaA_f]$. This

equation also includes a term representing the net loss of DnaA · ATP to the inactive pool $DnaA_i$. This net rate includes the reactivation of inactive DnaA forms to the active DnaA · ATP form (Sekimizu & Kornberg, 1988; Yung & Kornberg, 1988) as well as the inactivation processes. In addition to the above described production and degradation terms, the differential equation for $[DnaA_f]$ includes terms reflecting the mass balance laws for DnaA · ATP monomers due to binding to gene sites away from *oriC* and dilution effects. The differential equation for $DnaA_i$ is a simple linear equation with a term reflecting the net conversion of DnaA · ATP into the inactive pool and a term for dilution.

The next step of the mathematical model is to determine whether an initiation event occurs. We are assuming that DnaA binding to *oriC* is the rate limiting reaction in the initiation event. At the beginning of the computer simulation for the generation times examined, there are two *oriC*'s tagged zero or one. All descendants prior to division carry their ancestral *oriC* label. Whenever there are 30 or more molecules of DnaA · ATP bound to a particular *oriC*, an initiation event is assumed to occur, and replication proceeds until two new chromosomes are produced. The choice of 30 molecules is arbitrary, but necessarily fixed. This threshold value could be changed with appropriate changes in the kinetic rate constants. There is probably some critical number of DnaA molecules needed to affect a conformational change in the chromosomal structure which then allows entry of other proteins and subsequent priming of initiation of DNA replication (Bramhill & Kornberg, 1988).

Immediately after initiation, DnaA protein may dissociate from *oriC* because of the loss of superhelicity resulting from nicks in nascent DNA in the *oriC* domain. In the model, when an initiation event occurs, the DnaA · ATP molecules bound to the particular *oriC* are released into the cytoplasm, increasing the $[DnaA_f]$. A new *oriC* is created and both *oriC*'s enter a hemimethylated state for 8 min during which time binding of DnaA · ATP is prevented. We note that the model keeps track of both the age and the last occurring initiation event for a particular *oriC*. It is assumed that the two oldest *oriC*'s will separate at cell division; thus, all subsequent initiation events are identified and paired with one of these two *oriC*'s.

After testing for an initiation event, the mathematical model checks for a cell division. In order for a cell to divide, there must be two *oriC*'s which have an age greater than C + D, where the time required for elongation is assumed to be 45 min and the period of latency prior to cell division is taken to be 25 min. When this modeling condition for division is met, the cell divides with the volume split exactly in half. The concentrations are maintained, which means that the number of molecules of all of the relevant biochemical species are halved. The amount of DnaA · ATP bound to each *oriC* before cell division differs because of the stochastic process from which the number was obtained. This amount is maintained for each *oriC* in the newborn cell. As noted above for the initiation events, there are two labels associated with the two oldest *oriC*'s. At cell division only the *oriC*'s that are identified with the tag zero (possibly only one) are kept in the newborn cell, and these *oriC*'s are relabeled in an alternating pattern of zeroes and ones to identify which *oriC*'s will be kept in a subsequent cell division. At this point the cell cycle time τ is adjusted to reflect the time required for cell division, and this parameter

is then used in our subsequent calculations of mass of the cell. The final step in the model is to increase the time by 0·1 min and repeat the process unless the termination time is reached.

4. Simulation Results

In this section we discuss the results of our computer simulations of the mathematical model. Details of the equations and parameters used for the computer simulations are provided in the appendices. The computer simulations followed an exponentially growing cell for 500 min through its various initiations and divisions. The work was done on a PC-AT using STSC APL*PLUS^{TM} for the main program and GraphiC^{TM} for the graphic presentation.

From Table 2, we find the average times for cell division to be 44·8, 43·7 and 42·0 min for R values of 20, 10 and 5, respectively. The R value gives the number of random events for the protein DnaA · ATP (activator) to bind or dissociate to *oriC*, averaged over 0·01 min. These average cell division times are reasonably close to the experimental value, $\tau = 45$ min, from which we derived our parameters. The simulation with $R = 20$ had cell division times ranging from 37·6 to 50·7 min over the 11 cell cycles computed. The variance shows that most cell divisions occurred close to the average cell division time. For $R = 20$, there is a 95% confidence level for the cell divide times to occur in the interval $44{\cdot}8 \pm 6{\cdot}1$ min.

Figure 3 shows a series of graphs created by the mathematical model from the simulation with $R = 20$ for $\tau \cong 45$ min. The data used to generate the graphs are collected at 1 min intervals and, for clarity, only the first 300 minutes are shown. Figure 3(a) shows the total number of DnaA · ATP molecules that are free in the

TABLE 2

Summary of simulation data

Initial value for τ	τ (average doubling time, min)[a]	R[b]	σ^2(variance)[c]	No. of divisions	Initiation difference[d]
	44·8[e]	20	9·6	11	1·2
45 min	43·7	10	6·8	11	2·7
	42·0	5	5·8	11	2·9
	58·6	20	5·1	8	2·1
58 min	58·3[e]	10	12·2	8	2·8
	54·6	5	52·2	9	1·4
	42·0	20	5·7	11	1·4
38 min	39·3	10	6·1	12	1·4
	38·0[e]	5	15·1	13	2·2

[a] τ is the average doubling time computed from the data generated for a single cell traced over 500 min in the simulation.

[b] R is the number of random events averaged over each 0·01 min for the binding of DnaA · ATP to *oriC*.

[c] σ^2 is the variance for the average doubling time, τ.

[d] The initiation difference notes the average time between the first initiation event on one chromosome and the next initiation event on a second chromosome for a particular cell cycle.

[e] These simulations are used in Figs 3–6.

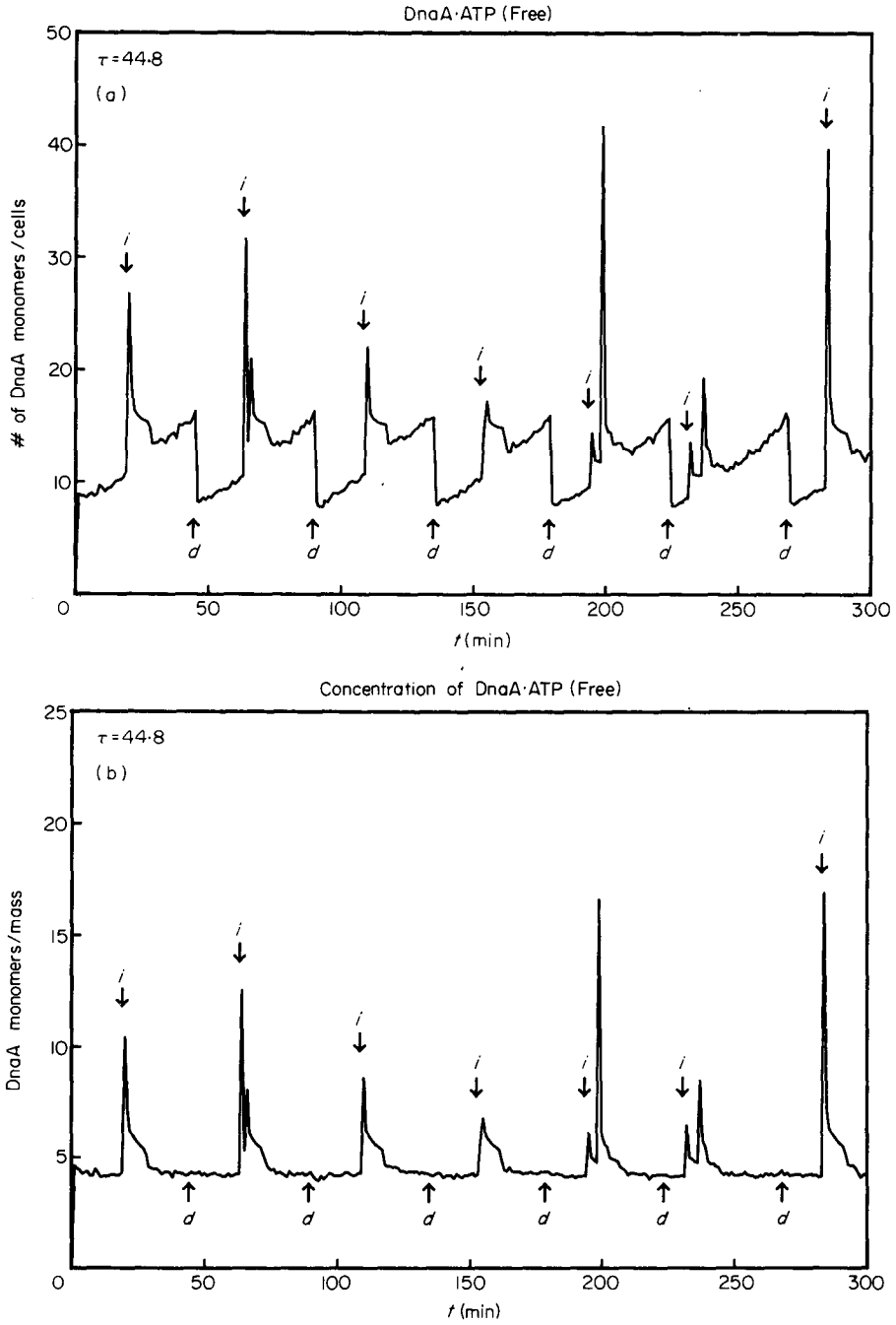

FIG. 3. (a). The number of DnaA · ATP (free) monomers/cell vs. time (min) for the generation time, $\tau = 45$ min; *i*, time of first initiation event; *d*, time of cell division. (b) The concentration of DnaA · ATP (free) monomers$\times 10^9/OD_{450}$ vs. time (min) for the generation time, $\tau = 45$ min; *i*, time of first initiation event; *d*, time of cell division. (c) The fraction of binding sites along the chromosome which are occupied by DnaA · ATP vs. time (min) for the generation time, $\tau = 45$ min; *i*, time of first initiation event; *d*, time of cell division. (d) The total number of DnaA monomers/cell vs. time (min) for the generation time, $\tau = 45$ min; *i*, time of first initiation event; *d*, time of cell division.

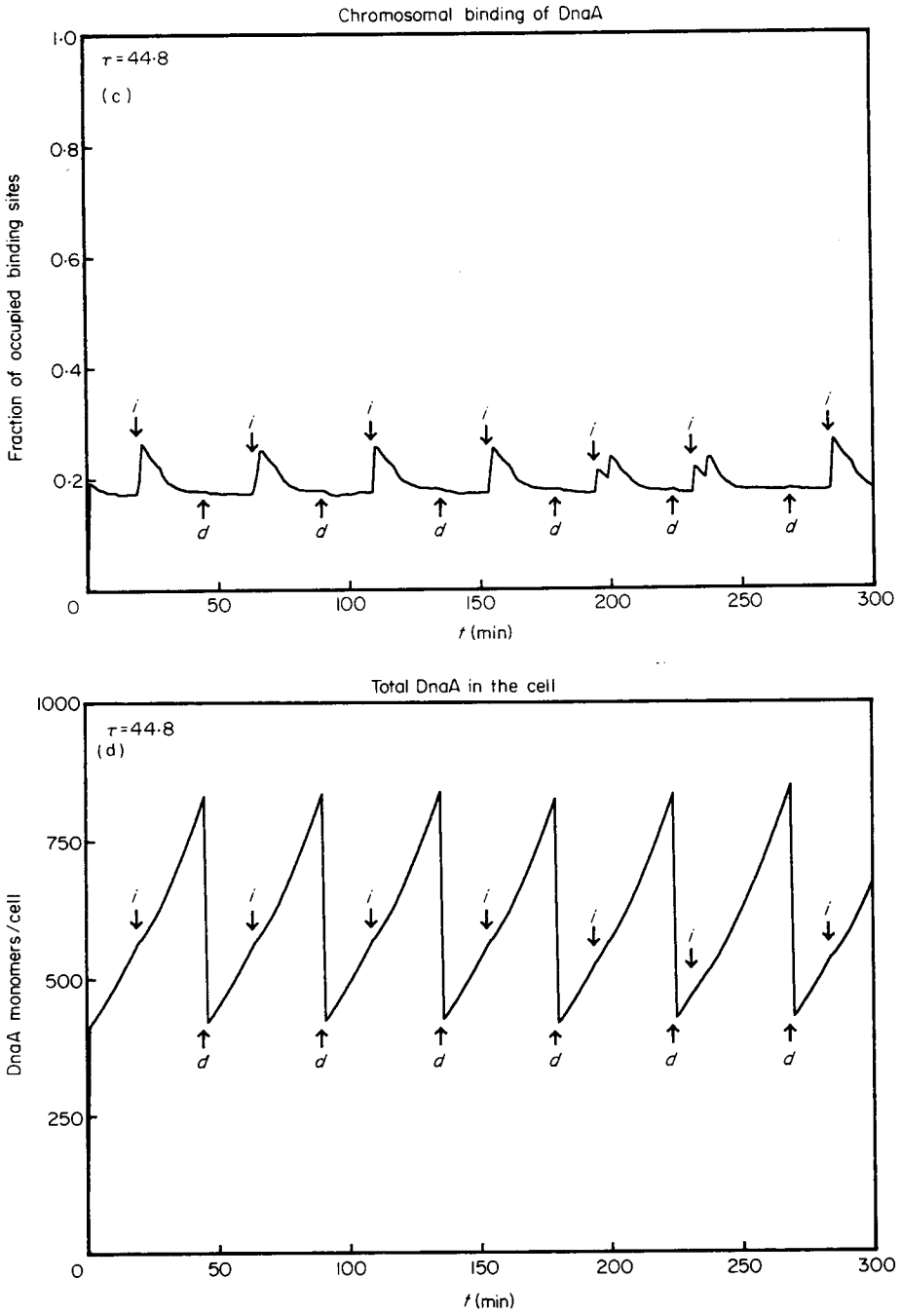

FIG. 3—continued

cytoplasm for binding or deactivation. The sharp peaks occur immediately following an initiation event and the release of the DnaA · ATP's bound to *oriC*. Doublets occur when the initiation on the other replication bubble is more than 1 min after the first initiation event, with the separation of the peaks showing the degree of syncrony of the two initiation events in the cell cycle. The height of the peaks is not particularly significant as free DnaA · ATP is rapidly absorbed on the chromosome or inactivated, meaning that the height is more a measure of how long after initiation the data is collected. Cell division times can be readily observed by the locations on the graph where the total number of DnaA · ATP monomers are halved. Figure 3(b) shows the concentration of the free DnaA · ATP in the cytoplasm as a function of time. The initiation events are very clear on this graph.

Figure 3(c) is presented to show how the DnaA · ATP released at initiation is rapidly absorbed on the chromosomal binding sites. The graph shows the fraction of sites on the chromosome occupied by the protein DnaA · ATP. Figure 3(d) shows the total number of DnaA monomers in the cell. Again the cell division times are clear as the number of DnaA monomers is halved. The values in this graph are in good agreement with the data in Table 1; Table 1 indicates there should be 407 DnaA monomers in the newborn cell and 554 DnaA monomers at initiation, values close to those in Fig. 3(d).

To test the validity of the model we need to determine if the kinetic constants used for $\tau = 45$ min will work for other growth rates. The only kinetic rates in the model which should vary with growth rate are the rate of production of DnaA · ATP from the unrepressed *dnaA* gene, k_{r1}, and the net rate of conversion of DnaA · ATP into the inactive pool, k_{r3}. Data of Chiaramello & Zyskind (1989) demonstrate growth rate regulation for the production of the protein DnaA. Thus, the ratios of the DnaA monomers/cell from Table 1 are used to provide a factor which multiplies the kinetic constant k_{r1}. We assume that the same factor applies to the kinetic constant k_{r3}, reflecting a similar increase (or decrease) in activity. As k_{r3} is a net flow from the active pool to the inactive pool, this rate change could be explained by several reasons. One reason is that the concentration of ATP does not increase with growth rate (Franzen & Binkley, 1961) which would limit the conversion of DnaA protein to the form complexed with ATP. Also, at slower growth rates the cell membrane to volume ratio is larger which would allow a greater acidic phospholipid-mediated conversion of DnaA · ADP to DnaA · ATP (Sekimizu & Kornberg, 1988). From Table 1, we find that the ratios of DnaA monomers at $\tau = 58$ min and $\tau = 38$ min to the DnaA monomers at $\tau = 45$ min are 0·4 and 1·38, respectively. The rate constants k_{r1} and k_{r3} are multiplied by 0·4 for $\tau = 58$ min and 1·38 for $\tau = 38$ min, while all other kinetic constants are kept the same as for the simulation with $\tau = 45$ min. The initial conditions were adjusted to reflect the experimental data in Table 1, and computer simulations of 500 min were run for the same R values as before. The results are given in Table 2 and illustrated in Figs 4 and 5. The average cell division times agree well with experimental times $\tau = 58$ min and $\tau = 38$ min, though the latter case demonstrated slightly longer cell divide times than would be expected for $R = 20$, yet still inside the 95% confidence interval. The graphs in Figs 4 and 5 are qualitatively similar to Figs 3(a) and 3(b) and can be

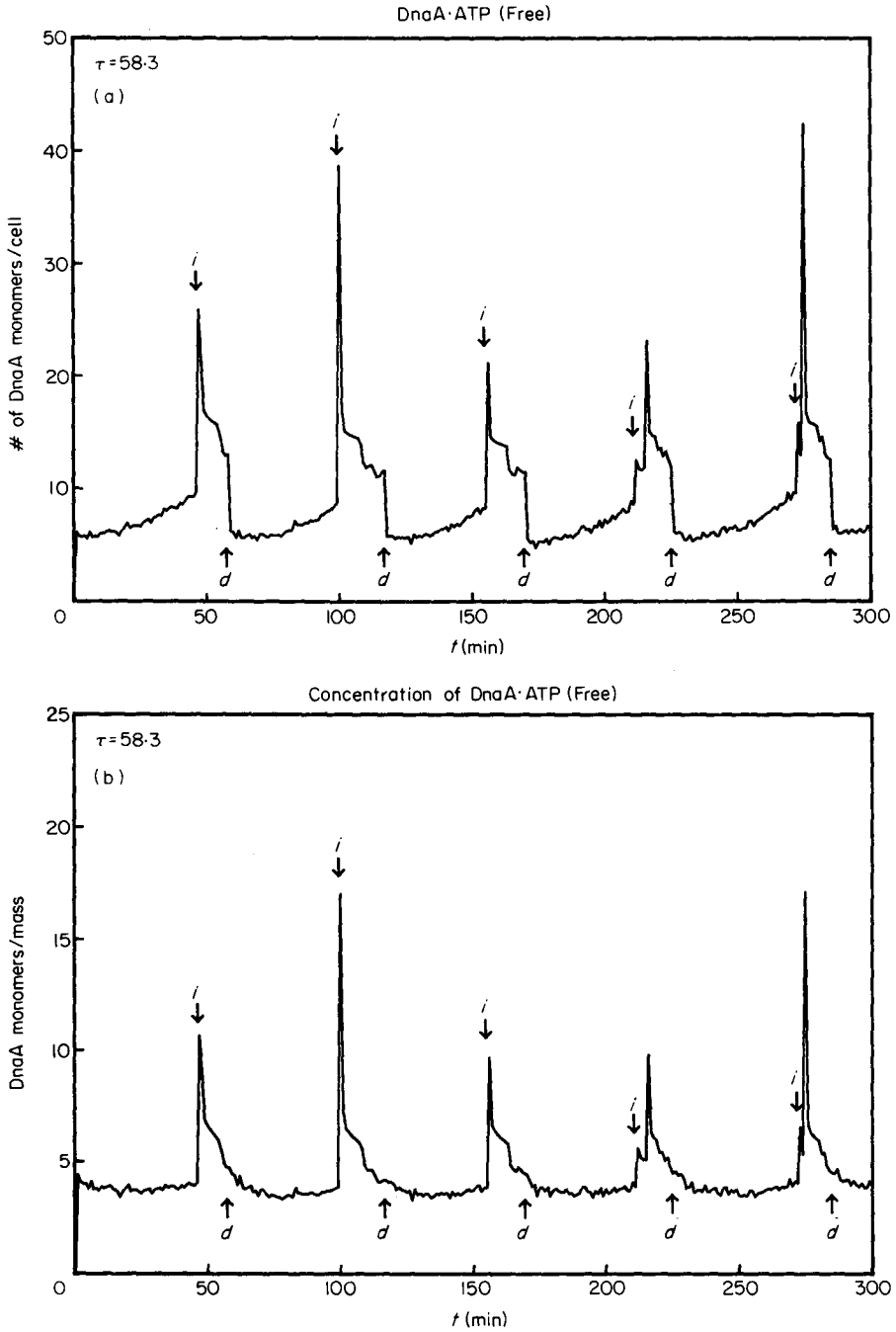

FIG. 4. (a). The number of DnaA · ATP (free) monomers/cell vs. time (min) for the generation time, $\tau = 58$ min; *i*, time of first initiation event; *d*, time of cell division. (b) The concentration of DnaA · ATP (free) monomers $\times 10^9/OD_{450}$ vs. time (min) for the generation time, $\tau = 58$ min; *i*, time of first initiation event; *d*, time of cell division.

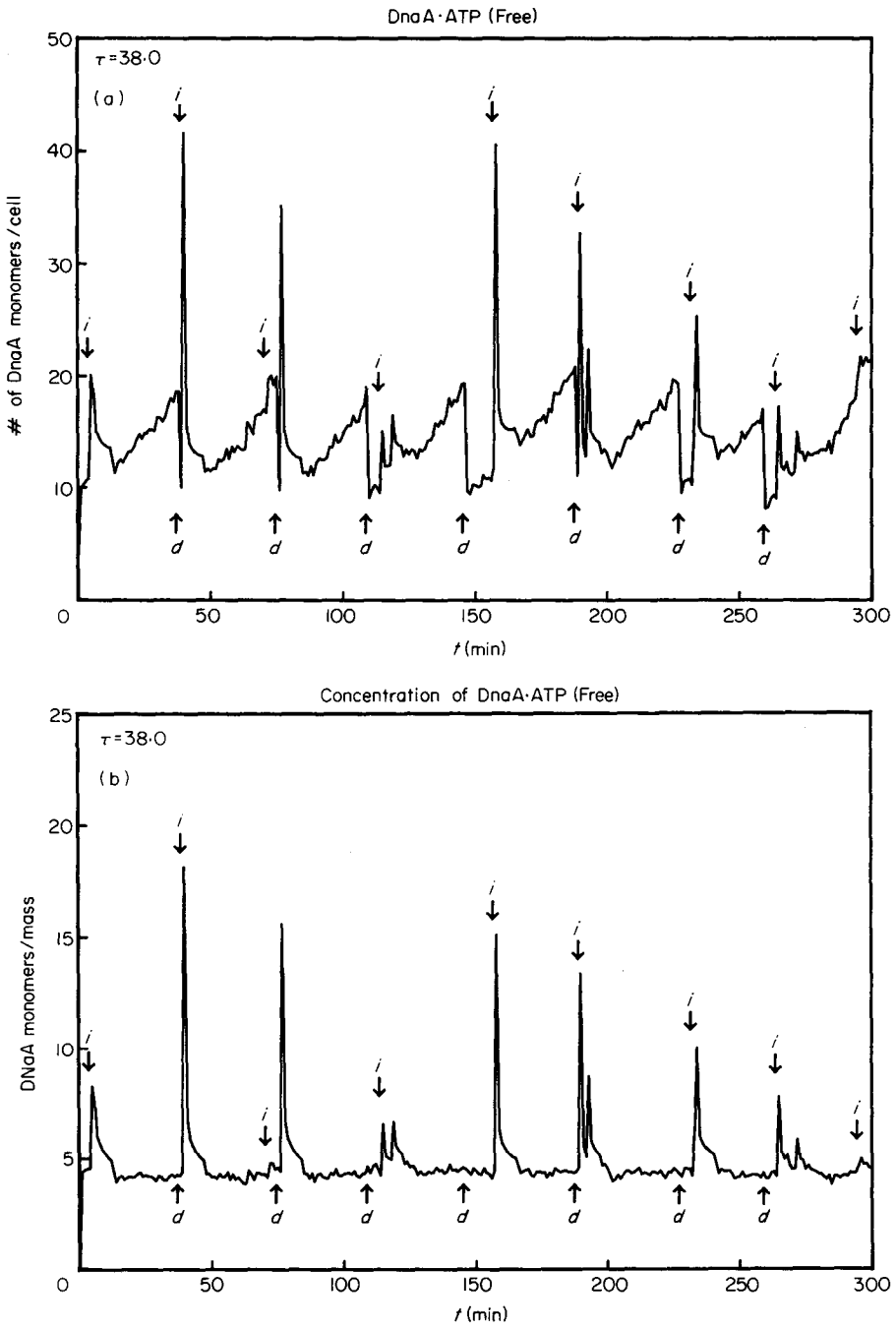

FIG. 5. (a). The number of DnaA · ATP (free) monomers/cell vs. time (min) for the generation time, $\tau = 38$ min; *i*, time of first initiation event; *d*, time of cell division. (b) The concentration of DnaA · ATP (free) monomers $\times 10^9/OD_{450}$ vs. time (min) for the generation ime, $\tau = 38$ min; *i*, time of first initiation event; *d*, time of cell division.

interpreted in a similar manner. Figure 5 is particularly interesting, as we observe two times ($t = 71{\cdot}1$ and $295{\cdot}3$ min) where a second initiation occurs during a cell cycle followed by no initiation during the next cell cycle. This event occurs as the initiation time and cell divide time are close for a cell that doubles every 38 min.

Figure 6 shows the concentrations for the total DnaA in the cells for three of the computer simulations. From Table 1 we find concentrations of 119, 215 and 255 DnaA ($\times 10^9$ monomers)/mass(OD_{450}) for cell doubling times $\tau = 58$, 45 and 38 min, respectively. The graph in Fig. 6 shows that the concentrations for the total DnaA in the cells remain relatively constant and fairly consistent with the experimental data. The values are slightly high for the simulations with $\tau = 58{\cdot}3$ and $44{\cdot}8$ min and low for the simulation with $\tau = 38{\cdot}0$ min.

The last column of Table 2 gives the average time between the first initiation event on one branch and the second initiation event on another branch. In a sense it is a measure of the degree of syncrony of the initiation events. With the exception of the entry for $\tau = 38$ min with $R = 5$, if this average initiation difference time is divided by the average doubling time, then we find the fraction of the time that the cell has three *oriC*'s rather than two or four. This ranged from 2·7% of the time for $\tau = 54{\cdot}6$, $R = 5$, to 6·9% of the time for $\tau = 42{\cdot}0$, $R = 5$. The special case $\tau = 38{\cdot}0$, $R = 5$, was shown to have three *oriC*'s 4·8% of the time and five *oriC*'s 1·0% of the time. If we assume a uniform distribution of cells, then this difference in the

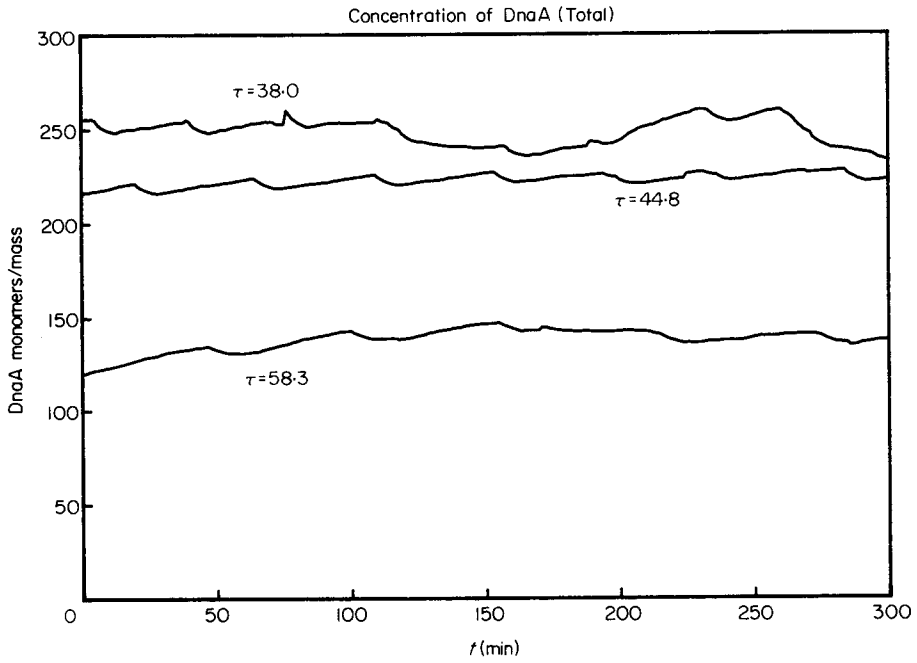

FIG. 6. Graphs of the concentrations of the total DnaA monomers $\times 10^9/OD_{450}$ vs. time (min) for the generation times, $\tau = 38$, 45 and 58 min.

timing of initiation events agrees very well with the results of Skarstad *et al.* (1986). Thus, the mathematical model gives the high degree of syncrony of initiation observed in experiments.

5. Discussion

We have developed a model that centers on the role of the DnaA protein as the principal control in the timing of initiation. The experimental results of Chiaramello & Zyskind (1989) presented in Table 1 and Fig. 2 show that the protein DnaA is growth rate regulated, which prevents its concentration from being the only factor governing the timing of initiation. However, the experimental findings discussed in the introduction suggest that reactions involving the protein, DnaA, with the origin of replication, *oriC*, may be rate limiting in the timing of initiation. Our model examines multiple roles for the protein, DnaA, based on known experimental results, to provide a plausible explanation for the controlling mechanism behind initiation of replication.

Figure 1 shows the key components of our model. Initiation occurs when a critical level of the activator, DnaA · ATP, accumulates on *oriC*. There is a reserve pool of DnaA · ATP available and probably bound to other sites along the chromosome. The DnaA protein is an autorepressor which implies that it maintains certain concentrations of free DnaA. Finally, the model considers a large inactive pool. Our model, which combines a stochastic binding to *oriC* with deterministic equations for the other reactions, was tested using computer simulations.

The mathematical model centers around the concentration of DnaA · ATP in the cytoplasm. From the data of Table 1 at slow growth rates we observe that there are not many more DnaA monomers/cell than we hypothesize are required for initiation, which indicates that the concentration of free DnaA · ATP is fairly low. Experimental evidence points to DnaA as the key control in the timing of initiation; however, Table 1 clearly shows that DnaA is growth rate regulated which seems to indicate that much of the DnaA at high growth rates is unable to activate *oriC*.

The computer simulations of our model show that the basal level of DnaA · ATP is maintained in a fairly narrow range of values which prevents the model cell from initiating too frequently. The small increase of the active form of DnaA at higher growth rates allows DnaA · ATP to be the key control in the timing of initiation as suggested by experimental evidence. In our model the large inactive pool of DnaA at higher growth rates can account for the data in Table 1 showing growth rate regulation of DnaA.

After an *oriC* initiation in the model, there is a release of its bound DnaA · ATP's into the cytoplasm resulting in sharp peaks in the graphs of the concentration of free DnaA · ATP. This rapid rise in concentration stimulates other *oriC*'s to initiate, resulting in a syncrony of initiation events. The model produces a syncrony of initiation events which agree closely with the experiments of Skarstad *et al.* (1988). Mahaffy (personal observation) found that when computer simulations were run with the DnaA protein released from *oriC* into the inactive pool rather than the active pool, the degree of syncrony decreased by a little more than a factor of two

(see the initiation difference column of Table 2). The *oriC* which has just initiated enters an eclipse period where DnaA . ATP cannot bind and can be explained experimentally by either an association of *oriC* with the membrane due to hemimethylation or the loss of supercoiling. This eclipse period in the model allows the concentration of the free DnaA · ATP to return to basal levels so that the development of the initiation complexes on the *oriC*'s proceed at a relatively steady rate and syncrony is not lost.

For the study presented in this paper we selected several parameters arbitrarily. As stated in the previous section, the only kinetic parameters in the model which would be affected by growth rate are k_{r1} and k_{r3}. By changing these values, the model successfully reproduced the cell cycle time and concentrations of DnaA observed in the experiments for different growth rates (see Table 2). More numerical studies are planned in the future to determine the sensitivity of the mathematical model to the kinetic parameters. In addition, some of the kinetic parameters in the model can be improved upon by further and more quantitative experimental results.

The modeling of this cellular control system suggests several important studies which should be investigated experimentally. The intracellular concentrations of the various forms of DnaA need to be known at different growth rates. Kinetic studies should be performed to determine rates of conversions among the various forms, such as hydrolysis of DnaA · ATP to DnaA · ADP and acidic phospholipid facilitated conversion of DnaA · ADP to DnaA · ATP. More information is needed on binding of DnaA to DNA. For example, the specific sequence requirements defining the DnaA box and the relevant binding constants should be determined.

Future models could consider competitive inhibition of binding to the consensus sequence in *oriC* and along the chromosome from binding of molecules in the inactive pool, $DnaA_i$. Yung & Kornberg (1989) have found that the DnaA · ATP · *oriC* complex is more stable than the DnaA · ADP · *oriC* complex when challenged with excess *oriC* DNA. Experiments could show that the cell membrane has a significant role in the initiation process in which case another model could examine the compartmentalization of the formation of the *oriC* · DnaA · ATP complex.

Our model should aid in prioritizing the choices of parameters that need to be determined experimentally and provide a framework for designing experiments to better understand how initiation of replication is controlled. Mathematically, we plan to investigate several of the qualitative properties of this distinctive dynamical system.

We would like to acknowledge Anne Chiaramello for her contributions to this paper, first, for her previously published data and, second, for her many insightful comments concerning the model described here.

REFERENCES

ATLUNG, T. (1984). *Molec. Gen. Genet.* **197,** 125.
ATLUNG, T., CLAUSEN, E. S. & HANSEN, F. G. (1985). *Molec. Gen. Genet.* **200,** 442.
ATLUNG, T., LØBNER-OLESEN, A. & HANSEN, F. G. (1987). *Molec. Gen. Genet.* **206,** 51.
BACHMANN, B. J. (1987). In: *Escherichia coli and Salmonella typhimurium cellular and molecular biology. Vol.* 2 (Neidhardt, F., ed.) p. 807. Washington, D.C.: ASM.

BRAMHILL, D. & KORNBERG, A. (1988). *Cell* **52,** 743.
BRAUN, R. E., O'DAY, K. & WRIGHT, A. (1985). *Cell* **40,** 159.
BREMER, H., CHURCHWARD, G. & YOUNG, R. (1979). *J. theor. biol.* **81,** 533.
BREMER, H. & DENNIS, P. P. (1987). In: *Escherichia coli and Salmonella typhimurium cellular and molecular biology. Vol.* 2 (Neidhardt, F., ed.) p. 1527. Washington, D.C.: ASM.
CHIARAMELLO, A. & ZYSKIND, J. W. (1989). *J. Bact.* **171,** 4272.
COOPER, S. & HELMSTETTER, C. E. (1968). *J. molec. Biol.* **31,** 519.
DONACHIE, W. D. (1968). *Nature, Lond.* **219,** 1077.
FRANZEN, J. S. & BINKLEY, S. B. (1961). *J. biol. Chem.* **236,** 515.
FULLER, R. S., FUNNEL, B. E. & KORNBERG, A. (1984). *Cell* **38,** 889.
FULLER, R. S. & KORNBERG, A. (1983). *Proc. natn. Acad. Sci. U.S.A.* **80,** 5817.
FUNNELL, B., BAKER, T. A. & KORNBERG, A. (1986). *J. biol. Chem.* **261,** 5616.
FUNNELL, B., BAKER, T. A. & KORNBERG, A. (1987). *J. biol. Chem.* **262,** 10327.
GOODWIN, B. C. (1965). *Adv. Enzyme Reg.* **3,** 425.
HANSEN, E. B., ATLUNG, T., HANSEN, F. G., SKOVGAARD, O. & VON MEYENBURG, K. (1984). *Molec. Gen. Genet.* **196,** 387.
HELMSTETTER, C. E. (1987). In: *Escherichia coli and Salmonella typhimurium cellular and molecular biology. Vol.* 2 (Neidhardt, F., ed.) p. 1594. Washington, D.C.: ASM.
INGRAHAM, J. L., MAALØE, O. & NEIDHARDT, F. C. (1983). In: *Growth of the Bacterial Cell,* p. 228. Sunderland, MA.: Sinauer Associates Inc.
INMAN, R. B. & SCHNÖS, M. (1987). *J. molec. Biol.* **193,** 377.
JACOB, R., BRENNER, S. & CUZIN, F. (1963). *Cold Spring Harb. Symp. Quant. Biol.* **28,** 329.
KOGOMA, T. & VON MEYENBERG, K. (1983). *EMBO J.* **2,** 463.
KOHARA, Y., AKIYAMA, K. & ISONO, K. (1987). *Cell* **50,** 495.
KÜCHERER, C., LOTHER, H., KÖLLING, R., SCHAUZU, M. A. & MESSER, W. (1986). *Molec. Gen. Genet.* **205,** 115.
LØBNER-OLESEN, A., SKARSTAD, K., HANSEN, F. G., VON MEYENBURG, K. & BOYE, E. (1989). *Cell* **57,** 881.
MATSUI, M., OKA, A., TAKANAMI, M., YASUDA, S. & HIROTA, Y. (1985). *J. molec. Biol.* **184,** 529.
MESSER, W., SEUFERT, W., SCHAEFER, C., GIELOW, A., HARTMAN, H. & WENDE, M. (1988). *Biochim. biophys. Acta* **951,** 351.
NOZAKI, N., OKAZAKI, T. & OGAWA, T. (1988). *J. biol. Chem.* **263,** 14176.
OGAWA, T., BAKER, T. A., VAN DER ENDE, A. & KORNBERG, A. (1985). *Proc. natn. Acad. Sci. U.S.A.* **82,** 3562.
OGDEN, G. B., PRATT, M. J. & SCHAECHTER, M. (1988). *Cell* **54,** 127.
OTHMER, H. G. (1976). *J. math. Biol.* **3,** 53.
POWELL, E. O. (1956). *J. gen. Microbiol.* **15,** 492.
SEKIMIZU, K., BRAMHILL, D. & KORNBERG, A. (1987). *Cell* **50,** 259.
SEKIMIZU, K. & KORNBERG, A. (1988). *J. biol. Chem.* **263,** 7131.
SEKIMIZU, K., YUNG, B. Y. & KORNBERG, A. (1988). *J. biol. Chem.* **263,** 7136.
SKARSTAD, K., BOYE, E. & STEEN, H. B. (1986). *EMBO J.* **5,** 1711.
SKARSTAD, K., VON MEYENBURG, K., HANSEN, F. G. & BOYE, E. (1988). *J. Bact.* **170,** 852.
SOMPAYRAC, L. & MAALØE, O. (1973). *Nature New. Biol.* **241,** 133.
WANG, Q. & KAGUNI, J. M. (1987). *Mol. Gen. Genet.* **209,** 518.
WOMBLE, D. D. & ROWND, R. H. (1987). *J. molec. Biol.* **195,** 99.
YUNG, B.-Y., M. & KORNBERG, A. (1988). *Proc. natn. Acad. Sci. U.S.A.* **85,** 7202.
YUNG, B.-Y., M. & KORNBERG, A. (1989). *J. biol. Chem.* **264,** 6146.
ZYSKIND, J. W., CLEARY, J. M., BRUSILOW, W. S. A., HARDING, N. E. & SMITH, D. W. (1983). *Proc. natn. Acad. Sci. U.S.A.* **80,** 1164.
ZYSKIND, J. W., DEEN, L. T. & SMITH, D. W. (1977). *J. Bacteriol.* **129,** 1466.

APPENDIX A

The Kinetic Equations

In this appendix we detail the kinetic equations used in the computer simulations. For an exponentially growing cell, if the initial mass following cell division is M_0

and τ represents the generation time, then the mass, $M(t)$, is given by

$$M(t) = M_0 2^{t/\tau},$$

where t is the time since the last cell division. By assuming that mass of the cell is proportional to the volume of the cell, we use $M(t)$ to compute the concentrations of the relevant biochemical species that are used in the binding and reaction calculations.

The probabilistic part of the model uses a Monte Carlo simulation scheme to determine the individual binding events of DnaA · ATP to *oriC*. We make the assumption that for a given Δt we can choose J such that only one binding event can occur during the time interval $\Delta t/J$. The probability that a molecule of DnaA · ATP binds to one of the four sites in *oriC* is given by

$$P_1(\Delta t/J) = \frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f],$$

where $[\mathrm{DnaA}_f]$ is the concentration of the free DnaA · ATP in the cytoplasm and k_0 is the binding constant for DnaA · ATP to *oriC* or to other DnaA molecules bound to *oriC*. The probability that no binding occurs is given by

$$P_0(\Delta t/J) = \left(1 - \frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f]\right).$$

The simulations are for the interval Δt. With the particular values of k_0 and J used and the range over which $[\mathrm{DnaA}_f]$ varied, the probability of more than two binding events occurring during the interval Δt is very small. This allows us to ignore the higher order terms in the expansions below, resulting in the following:

$$\begin{aligned}
P_0(\Delta t) &= \left(1 - \frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f]\right)^J \\
&\cong 1 - \Delta t 4k_0[\mathrm{DnaA}_f] + \frac{(J-1)\Delta t^2}{2J}\left(4k_0[\mathrm{DnaA}_f]\right)^2, \\
P_1(\Delta t) &= \binom{J}{1}\left(1 - \frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f]\right)^{J-1}\left(\frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f]\right) \\
&\cong \Delta t 4k_0[\mathrm{DnaA}_f] - \frac{(J-1)\Delta t^2}{J}(4k_0[\mathrm{DnaA}_f])^2, \\
P_2(\Delta t) &= \binom{J}{2}\left(1 - \frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f]\right)^{J-2}\left(\frac{\Delta t}{J} 4k_0[\mathrm{DnaA}_f]\right)^2 \\
&\cong \frac{(J-1)\Delta t^2}{2J}(4k_0[\mathrm{DnaA}_f])^2,
\end{aligned} \tag{1}$$

where $P_i(\Delta t)$ is the probability that i molecules of DnaA bind to *oriC* over the interval Δt.

The model also allows for dissociation of at most one molecule over the interval Δt. We have assumed that only the upper most molecules in our binding scheme can become unbound. Let the probability that a DnaA · ATP monomer becomes unbound be $P_{-1}(\Delta t) = \Delta t k_d N$, where k_d is the kinetic constant for the reverse reaction to the binding of DnaA · ATP to *oriC*, and N is the number of exposed molecules which, by assumption, is four unless there are fewer molecules bound to *oriC*, in which case it is the total number of DnaA molecules bound to *oriC*. Combining this information with eqn (1), we find the resultant probabilities used in the Monte Carlo part of our simulation. If we denote $Q_i(\Delta t)$ to be the net probability that i molecules are added to *oriC* over the interval Δt, then we have the following:

$$\begin{aligned} Q_2(\Delta t) &= P_2(\Delta t)(1-\Delta t k_d N), \\ Q_1(\Delta t) &= P_1(\Delta t)+\Delta t k_d N(P_2(\Delta t)-P_1(\Delta t)), \\ Q_0(\Delta t) &= P_0(\Delta t)+\Delta t k_d N(P_1(\Delta t)-P_0(\Delta t)), \\ Q_{-1}(\Delta t) &= \Delta t k_d N P_0(\Delta t). \end{aligned} \tag{2}$$

The probabilities in eqn (2) are used to partition a uniform distribution on the interval [0, 1] into four sets which relate to whether molecules for a particular *oriC* are added or subtracted. A random number in this interval is chosen, and depending from which set it comes, we take a molecule from the ones bound to *oriC* (Q_{-1}), add one or two molecules (Q_1 or Q_2), or leave *oriC* in its current state (Q_0). This completes the mathematical description of the stochastic part of the model used in section 3.

The first part of the deterministic phase of the model is the computation of the total number of gene sites, *GS*, which have the consensus sequence. As stated in section 3, these binding sites are assumed to be added regularly to all actively elongating chromosomes, AC, which correspond to the number of *oriC*'s with an age between 0 and C. From this information we find that

$$GS(t+\Delta t) = GS(t) + AC k_{gs} \Delta t,$$

where k_{gs} is the rate of formation of gene sites per elongating chromosome.

From section 3 we find the non-linear differential equations which govern the concentrations of DnaA_g, DnaA_f and DnaA_i. The concentration of DnaA · ATP bound to the gene sites is determined by the equation:

$$\frac{\mathrm{d}[\mathrm{DnaA}_g]}{\mathrm{d}t} = k_{g1}[\mathrm{DnaA}_f]([GS]-[\mathrm{DnaA}_g]) - (k_{g2}+\mu)[\mathrm{DnaA}_g], \tag{3}$$

where k_{g1} is rate of formation of bonds between DnaA · ATP in the cytoplasm with unbound gene sites, k_{g2} is the rate of dissociation of the bound form, and $\mu = \ln(2)/\tau$ is the dilution rate with τ the cell doubling time.

The equation for DnaA · ATP in the cytoplasm is

$$\begin{aligned} \frac{\mathrm{d}[\mathrm{DnaA}_f]}{\mathrm{d}t} = {} & \frac{k_{r1}}{1+K_{r2}([\mathrm{DnaA}_f]+\gamma[\mathrm{DnaA}_i])} - k_{g1}[\mathrm{DnaA}_f]([GS]-[\mathrm{DnaA}_g]) \\ & + k_{g2}[\mathrm{DnaA}_g] - (k_{r3}+\mu)[\mathrm{DnaA}_f]. \end{aligned} \tag{4}$$

The first expression on the right hand side of eqn (4) represents the production of $DnaA_f$ using standard biochemical kinetics for end product repression. Using the experimental evidence given in the introduction, we assume that one molecule of DnaA binds to the operator region of *dnaA* to create the repression complex. The repressor concentration is determined from the concentration of $DnaA_f$ and from certain forms of DnaA that are in the inactive state, $[DnaA_i]$. For modeling purposes we assume that the amount of $DnaA_i$ which is available for repression is a fraction of the total $DnaA_i$. One would also predict that the different forms of $DnaA_i$, such as DnaA unbound to a nucleotide or DnaA · ADP, would have binding rates different from $DnaA_f$. Thus, the constant γ reflects both the fraction of $DnaA_i$ which is available for repression and possible differences in the effective binding rates of these relevant repressor forms compared to $DnaA_f$. The constant K_{r2} is an equilibrium constant for the affinity of DnaA · ATP with the operator region of the *dnaA* gene. The rate constant k_{r1} reflects the transcription and translation rates for producing DnaA protein from the *dnaA* gene. The terms with k_{g1} and k_{g2} reflect the balance laws for binding of DnaA · ATP to the chromosome. Finally, k_{r3} is the rate constant for conversion of $DnaA_f$ to $DnaA_i$ as described in section 3. Again, μ is the rate constant for dilution effects. Note that binding of DnaA · ATP to *oriC* is computed in the probabilistic part of the model, its effect has already been included and, thus, does not appear in the differential equation.

The differential equation for the concentration of inactive DnaA, $[DnaA_i]$, is given by:

$$\frac{d[DnaA_i]}{dt} = k_{r3}[DnaA_f] - \mu[DnaA_i]. \tag{5}$$

This completes the deterministic portion of the model. A single step numerical integration using a fourth order Runge–Kutta scheme is employed to update the concentrations given in eqns (3), (4) and (5).

Section 3 describes the tests performed to determine if initiation or division occurs. Every 0·1 min the simulation cycles sequentially through the stochastic and deterministic parts of the mathematical model. At each minute in the simulation the computer program stores all relevant data for the plots produced in section 4.

APPENDIX B

The Kinetic Parameters

From Appendix A it is apparent that there are many kinetic parameters which must be chosen. In this appendix we discuss what values were taken and how we chose the particular values. The data of Chiaramello & Zyskind (1989) were used for the growth rate of $\tau = 45$ min to begin our search for parameter values. For this growth rate, the mass of the newborn cell, M_0, was found to be $1{\cdot}89 \times 10^{-9}$ OD_{450}. Recall that both M_0 and τ are recomputed when a cell division occurs in the program, so the experimental data are used only to initiate the program.

As stated above, we arbitrarily chose the value of 30 molecules of DnaA · ATP bound to *oriC* to determine when an initiation event occurred. Using a cell which

is dividing every 45 min and noting the C+D times given before, we see that an initiation event occurs approximately 20 min after cell division, with the two *oriC*'s replicating to become four. Our model assumes that the first 8 min after initiation do not allow binding. This leaves 37 min for the cell to accumulate 30 molecules of DnaA · ATP per *oriC*, or 0·8 molecules/min are added on average. If DnaA · ATP molecules are added at regular intervals, then *oriC*'s at the time of cell division must have about 13 molecules of DnaA · ATP bound to them. Thus, the initial state of the *oriC* origins in our simulation for $\tau = 45$ is taken to have two *oriC*'s each with an age of 25 min, a time of 25 min since the last initiation event, and a binding state of 13 molecules of DnaA · ATP.

The computer simulation needs numerical values for the kinetic constants k_0 and k_d, which are assumed to be average binding and dissociation constants for the stochastic part of the model. Experimental studies suggest that the binding of DnaA to *oriC* is very stable (Yung & Kornberg, 1989). With this information we assumed a binding half-life of about 2 min, which implies $k_d \cong \ln(2)/2 \cong 0{\cdot}35$ (min^{-1}). We arbitrarily chose the initial cell to have eight molecules in the cytoplasm which gives $[\mathrm{DnaA}_f] \cong 4{\cdot}2$ ($\times 10^9$ molecules/OD_{450}). As 0·8 (molecules/min) $\cong 4k_0\,[\mathrm{DnaA}_f] - 4k_d$, we can compute,

$$k_0 = 0{\cdot}13\ (\times 10^{-9}\ OD_{450}/\text{molecule-min}).$$

More experimental results are needed to improve the selection of the rate constants discussed above. As noted in section 3, the kinetic rates k_0 and k_d are taken to be average constants for the overall rates of binding and dissociation; however, it is clear that they should depend on the number of DnaA · ATP molecules bound to *oriC*.

As there are approximately 10^7 nucleotides (5×10^6 bp) in the chromosome of *E. coli* (Kohara *et al.*, 1987), we would predict about $10^7/4^7 2^2 \cong 150$ consensus sequences along the chromosome. This gives $k_{gs} = (150/45) = 3{\cdot}33$ binding sites/min for the rate at which active gene sites are formed along an elongating chromosome.

The low $[\mathrm{DnaA}_f]$ which is used in these parameter calculations assumes that most of the DnaA · ATP would be bound to either the *oriC*'s or to the other specific binding sites along the chromosome. As there is a sharp rise in the $[\mathrm{DnaA}_f]$ after an initiation event, the average $[\mathrm{DnaA}_f]$ is higher than the value used to compute k_0. We chose an average $[\mathrm{DnaA}_f]$ of 4·7 ($\times 10^9$ molecules/OD_{450}) for the subsequent quasi-steady state calculations. Equation (3) is used to determine how many DnaA · ATP molecules are bound to the chromosome away from *oriC*. It is predicted that DnaA would bind to the different forms of the consensus sequence at different rates; however, we are assuming that the kinetic constants k_{g1} and k_{g2} are average binding rates. We arbitrarily chose the binding half-life along the chromosome to be 2 min as before, which gives $k_{g2} = 0{\cdot}35$ (min^{-1}). From the quasi-steady state calculation we find

$$k_{g1} = \frac{k_{g2}[\mathrm{DnaA}_g]}{[\mathrm{DnaA}_f]([GS] - [\mathrm{DnaA}_g])} \cong 0{\cdot}0186\ (\times 10^{-9}\ \mathrm{OD}_{450}/\text{molecule-min}),$$

assuming 20% of the binding sites are occupied. The newborn cell has one complete chromosome and one which has been elongating for 25 min, as the initiation event

occurs at $t = 20$ in the cell cycle. This implies that the amount of active chromosome in a newborn cell is $1+25/45 \cong 1{\cdot}56$ active chromosomes. The initial number of gene sites occupied is found by taking 150 (gene sites/chromosome) $\times$ 1·56 (chromosome/cell) $\times$ 0·2 $\cong$ 47 (sites bound/cell).

The last collection of kinetic constants which must be computed are in eqn (4). We arbitrarily chose $K_{r2} = 2$ ($\times 10^{-9}$ OD_{450}/molecules), which represents a moderately strong degree of repression, and $\gamma = 0{\cdot}01$, which is equivalent to a 1% contribution from $DnaA_i$. From the data in section 2 for $\tau = 45$ min, we see that there are 408 molecules of DnaA for the average newborn cell. This implies that the cell must produce 9·1 molecules of DnaA/min. The initial state given above has 26 molecules of DnaA · ATP bound to *oriC* and 47 molecules of DnaA · ATP bound to the chromosome. The number of DnaA · ATP molecules in the cytoplasm of the newborn cell is eight, as described above, which gives the number of $DnaA_i$ molecules as 327. The data show that the average cell mass is 2·7 ($\times 10^{-9}$ OD_{450}). Combining these results, we find the following:

$$k_{r1} = \frac{9{\cdot}1}{2{\cdot}7}[1 + K_{r2}([DnaA_f] + \gamma[DnaA_i])] \cong 46{\cdot}7(\times 10^9 \text{ molecules}/OD_{450} \cdot \text{min}).$$

The growing cell causes a loss of $2{\cdot}7[DnaA_f]\ln(2)/\tau \cong 0{\cdot}2$ ($\times 10^9$ molecules/OD_{450} · min) due to dilution. Using the initial conditions for the DnaA · ATP bound to *oriC* and the chromosome, we see that over one cell cycle 1·6 molecules of DnaA · ATP/min must be added. Combining the production rate with the losses due to dilution and binding at $\tau = 45$ min, we observe a net loss of 7·3 molecules of DnaA · ATP/min into the inactive pool, $DnaA_i$. This implies that

$$k_{r3} \cong \frac{7{\cdot}3}{2{\cdot}7[DnaA_f]} \cong 0{\cdot}58(\text{min}^{-1}),$$

and concludes the computation of the kinetic constants used for the simulation at $\tau = 45$ min.

APPENDIX C

Summary of Variables and Kinetic Parameters

Below is a list of the variables and kinetic parameters used in the mathematical model:

AC	number of actively elongating chromosomes.
$[DnaA_f]$	concentration of DnaA · ATP in the cytoplasm free to undergo reactions.
$[DnaA_g]$	concentration of DnaA · ATP bound to the consensus sequence.
$[DnaA_i]$	concentration of various inactive forms of DnaA protein.
γ	constant representing the binding strength and the fraction of $DnaA_i$ associated with the autorepression of the *dnaA* gene.
GS	number of gene sites with the consensus sequence.

k_0	rate of binding DnaA · ATP to *oriC*.
k_d	dissociation constant for DnaA · ATP from *oriC*.
k_{gs}	rate of formation of the consensus sequence along an elongating chromosome.
k_{g1}	rate of binding of DnaA · ATP to the consensus sequence.
k_{g2}	rate of dissociation of DnaA · ATP to the consensus sequence.
k_{r1}	rate of production of DnaA · ATP from the unrepressed *dnaA* gene.
K_{r2}	equilibrium constant for association of DnaA · ATP with the operator region of the *dnaA* gene.
k_{r3}	net rate of conversion of DnaA · ATP into the inactive pool, $DnaA_i$.
M_0	initial mass following a cell division.
$M(t)$	mass of the cell as a function of time.
μ	dilution rate which equals $\ln(2)/\tau$.
τ	generation time.